Accepted author's manuscript. Published in final edited form as: Osteoporosis International 2020 epub. Publisher DOI: 10.1007/s00198-020-05405-0

Reference Values and Clinical Predictors of Bone Strength for HR-pQCT-Based Distal Radius and Tibia Strength Assessments in Women and Men

Authors:

Anna K. Stuck [1)], Denis Schenk [2)], Philippe Zysset [2)], Lukas Bütikofer[3)], Andrea Mathis[2)], Kurt Lippuner[5)]

Author's affiliations:

1) Department of Geriatrics, Inselspital, Bern University Hospital, and University of Bern, Bern, Switzerland.
2) ARTORG Center for Biomedical Engineering Research, University of Bern, Freiburgstrasse 3, 3010 Bern, Switzerland.
3) CTU Bern, University of Bern, 3010 Bern, Switzerland.
4) Department of Osteoporosis, Bern University Hospital, University of Bern, Bern, Switzerland.

Keywords:

Reference values, HR-pQCT, bone strength, failure load, finite element analysis, multiple stacks, clinical predictors

Total Word Count: 3948

Abstract Word Count: 250

Conflict of interest: Anna K. Stuck, Denis Schenk, Philippe Zysset, Lukas Bütikofer, Andrea Mathis and Kurt Lippuner declare that they have no conflict of interest.

Acknowledgements: The authors would like to thank all the participants who graciously devoted time to participate in the study. We thank Furkan Gazozcu for scan acquisition and participant recruitment and collaborators of the University Clinic of Osteoporosis at the University Hospital Bern for DXA measurements and performance of quality protocols

Funding: This work was in part supported by the „Forschungsfonds der Geriatrischen Universitätsklinik", Bern/Switzerland. The funder had no role in, study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Corresponding Author:

Prof. Dr. med. Kurt Lippuner

Universitätspoliklinik für Osteoporose

Inselspital

CH-3010 Bern

Switzerland

Email: kurt.lippuner@insel.ch

Phone: +41 31 632 31 28

Mini-Abstract

Reference values for radius and tibia strength using multiple stacks HR-pQCT with homogenized finite element analysis analysis are presented in order to derive critical values improving risk prediction models of osteoporosis. . Gender and neck hip aBMD were independent predictors of bone strength.

Abstract

Purpose

The purpose was to obtain reference values for radius and tibia bone strength computed by homogenized finite element analysis (hFE) using multiple stacks with a high-resolution peripheral quantitative computed tomography (HR-pQCT).

Methods

Male and female healthy participants aged 20-39 years were recruited at the University Hospital Bern. They underwent interview, clinical examination including hand grip, gait speed, and DXA of the hip. The nondominant forearm and tibia were scanned with a double and a triple stack protocol, respectively, using HR-pQCT (XCT II, SCANCO Medical AG). Bone strength was estimated by hFE analysis and reference values were calculated using quantile regression. Multivariable analyses were performed to identify clinical predictors of bone strength.

Results

Overall, 46 women and 41 men were recruited with mean ages of 25.1 (sd 5.0) and 26.2 (sd 5.2) years. Sex-specific reference values for bone strength were established. Men had significantly higher strength for radius (mean (sd) 6642 (1797) N vs. 4107 (1199) N; $p<0.001$) and tibia (18200 (4219) N vs. 11971 (3153) N; $p<0.001$) than women. In the two multivariable regression models with and without total hip areal bone mineral density (aBMD), the addition of neck hip aBMD significantly improved the model ($p<0.001$). No clinical predictors of bone strength other than gender and aBMD were identified.

Conclusion

Reference values for radius and tibia strength using multiple HR-pQCT stacks with hFE analysis are presented and provide the basis to help refining accurate risk prediction models. Neck hip aBMD and gender were significant predictors of bone strength.

Introduction

Osteoporosis is a prevalent disease with major impact on health outcomes and functionality of primarily affected older people. Recent evidence shows, that osteoporosis and osteopenia in the distal forearm predict all-cause mortality [1]. Currently, the diagnosis of osteoporosis is based on areal bone mineral density (aBMD) measured by dual-energy X-ray absorptiometry (DXA). However, DXA has several limitations including that it is only a surrogate of bone strength and that it does not isolate the respective contributions of cortical bone geometry and trabecular bone microstructure to bone strength. Thus, half of all incident fractures occur in postmenopausal women who were not diagnosed with osteoporosis, i.e. in women with an aBMD T-score above the -2.5 threshold defined by the WHO as an operational definition for osteoporosis [2, 3].

Second generation high-resolution peripheral quantitative computed tomography (HR-pQCT) allows to assess bone architecture including microstructural parameters with an unprecedented resolution of 61 microns [4]. Bone stiffness and strength estimated by HR-pQCT-based finite element analysis have been shown to be associated with prevalent fractures [4-7]. In addition, Samelson et al. [8] showed in their comprehensive prospective cohort that cortical and trabecular bone microarchitecture were independent predictors of incident fracture risk in older women and men. HR-pQCT measurements at the distal tibia were even correlated with lumbar vertebra failure loads [9]. HR-pQCT may be helpful as an add-on diagnostic tool to approach specific diagnostic or treatment challenges in clearly defined populations, such as patients with hyperparathyroidism, diabetes mellitus, renal osteodystrophy, or haematological malignancies.

Up to date, major cohort studies indicating reference values for compressive bone strength were conducted using micro finite element analysis (µFE) based on single stack HR-pQCT measurements [5, 10-12]. Unfortunately, the “weakest” and most predestinated location for fracture (i.e. location of Colles’ fracture for radius [13] and metaphyseal area for tibia) is not entirely contained in single stacks HR-pQCT scans. Mueller et al. and Varga et al. demonstrated, that the most relevant region to determine radius failure load was located just below the end of the subchondral plate [14, 15]. Thus, bone strength may be overestimated using single stacks and patients at risk for fractures may be missed. Alternative HR-pQCT measurements using multiple stacks (Figure 1) have been combined with the more efficient homogenized finite element analysis (hFE) to address this limitation [4]. The combination of multiple stacks with hFE is a promising approach for transition of HR-pQCT into clinical practice. However, so far no reference data are available on this new methodology combining multiple stacks with hFE.

The primary objective of the study was to obtain reference values for radius and tibia compressive bone strength in the young healthy male and female Swiss population computed by hFE analysis using multiple stacks with a second-generation HR-pQCT (XCT II, SCANCO Medical AG). In addition, clinical predictors of bone strength were evaluated.

Methods

Participants

We recruited male and female participants aged 20-39 years through local advertisement at the University of Bern and the University Hospital Bern between August and December 2018. Participants included in this study were healthy individuals, while those suffering from any medical conditions or taking any medications known to affect bone metabolism were excluded. The present study was conducted with the approval of the Swiss Ethics committee. All participants provided written informed consent.

Data collection

Participants were invited to attend a single clinical visit at the study site of the University Policlinic for Osteoporosis at the University Hospital of Bern. At this visit an interview, clinical assessment, DXA and HR-pQCT measurements were performed. Based on standardized questionnaire data on demographics, medical and fracture history, medication, risk factors for osteoporosis, and dietary habits were collected. Height and weight were measured on site. For evaluation of physical performance, the Short Physical Performance Battery (SPPB)[16] including gait speed and repeated chair rising test were used. Grip strength of both hands was measured three times with a dynamometer (Takei Physical Fitness Test, Japan) on a scale from 0 to 100 kg. Thereafter, participants underwent a standard DXA examination of the hip according to the usual protocol of the University Policlinic of Osteoporosis. Machine calibration, daily and weekly quality assurance assessments were performed and monitored according to guidelines of manufacturer. Results were expressed in absolute values of aBMD in g/cm^2 and as T- and Z-scores at the femoral neck and total hip. Study data were collected and managed using REDCap electronic data capture tools hosted at Clinical Trial Unit at the University of Bern. REDCap (Research Electronic Data Capture) is a secure, web-based software platform designed to support data capture for clinical research studies.

HR-pQCT

The nondominant forearm (or the nonfractured forearm in the case of a previous wrist fracture) and tibia were scanned using HR-pQCT (XtremeCT II, Scanco Medical, Brütisellen Switzerland) with a resolution of 61 µm and following the standard human *in vivo* scanning protocol (60 kVp, 1000 mA, 100-ms integration time). Double stack scans (336 slices) were performed at the distal radius and triple stack scans (504 slices) at the distal tibia. One stack covers a scan region of 10.2 mm. The positions of the two stacks covering the regions of interest in the radius are shown in Figure 1 (Panel A). The new reference line for radius images is defined according to Bonaretti et al. [17]. The positions of the three stacks covering the regions of interest in the tibia are shown in Figure 1 (Panel B). The reference line

at the tibia corresponds to the recommendation of the manufacturer (SCANCO Medical AG, Brütisellen, Switzerland). The region of interest covers the metaphyseal location of distal tibia fractures and accounts for the larger bone size of the tibia compared to the radius.

Quality of all HR-pQCT images were scored for motion artefacts according to a motion scale of 1 (no motion) to 5 (significant blurring of the periosteal surface, discontinuities in the cortical shell, or streaking in the soft tissue) according to a previously described method. [18] The interrater agreement between the two reviewers for quality grading was substantial (*k* value 0.73). Scans scored as 4 and 5 on the motion scale were considered as non-evaluable, and were repeated hereinafter up to three times for both tibia and radius until a quality score of scans of 1-3 was ascertained.

Prior to scan acquisition, the subject's arm or leg was immobilized in an anatomically formed brace provided by the manufacturer. Quality control was monitored using the calibration phantom provided by the manufacturer, using the daily and weekly protocols within the scanner's software. The HR-pQCT images were analysed using the standard manufacturer's method. An automated contouring scheme was used to segment the periosteal surface followed by a threshold-based algorithm that separates the cortical and trabecular bone compartments.

The primary HR-pQCT outcome variables were hFE estimated failure loads for radius resp. tibia; key predefined secondary outcome variables were total volumetric bone mineral density (vBMD), trabecular vBMD, trabecular bone area, trabecular thickness, cortical vBMD, cortical bone area, cortical perimeter, cortical thickness, and cortical porosity.

Homogenized finite element analysis

The hFE method used to calculate bone stiffness and strength was described in detail by Hosseini et al. [4] and Arias-Moreno et al. [19]. Briefly, all HR-pQCT images, independently of their grading score, were analyzed using the standard clinical image processing and FE workflow implemented on the scanner software (IPL V5.16/FE-v02.02, Scanco Medical AG). This included automatic definition of the periosteal, cortical and trabecular contour and generation and evaluation of the hFE models. Cortical and trabecular compartments were segmented using Gauss filtering (sigma = 0.8 / support = 1 voxel) and fixed thresholding (trabecular bone: 320mgHA/cm3, cortical bone: 450 mgHA/cm3). The nonlinear hFE models were generated based on the downscaled periosteal mask using 8-node brick elements with an edge length of 1.7mm (downscale factor 28). Homogenized properties were then assigned to each element based on the masks, the segmentations, BMD and the mean interception length (MIL [20]) fabric tensor. Displacements of nodes at the proximal surface were suppressed in all directions, those at the distal surface were only suppressed in the two in-plane directions. The models were compressed perpendicular to the surface until failure and reaction force and displacement of the top

surface were recorded. Stiffness was computed from the first load step (force/displacement) and strength was defined as the maximal recorded reaction force.

Long-term calibration correction

During the study, we observed a drift in mean volumetric bone mineral density (vBMD) of the daily quality control measurements QC1. According to the manufacturer's the recommendations a recalibration of the BMD calibration curve was performed after the drift exceeded +1% of the phantom norm BMD value. We fitted a curve through each section to specify the difference in vBMD for each day we performed measurements. The fitting filters the daily fluctuations of individual quality control measurements and thus represents the long-term calibration drift. To investigate the influence of this BMD drift on hFE strength, we changed the BMD calibration curve of a subgroup of 6 different radius measurements to the maximum permissible calibration drift of +-1%. These images where then processed as described above and hFE strength was compared to the initial value by means of linear regression. For each measurement we could then individually define the long-term calibration error in % BMD at the day of measurement and compensate the drift by adjusting hFE strength according to the linear regression.

Statistical analysis

Sample size simulations using quantile regression with 95% bootstrap confidence intervals on samples simulated from a reference normal distribution with mean and standard deviation according to references of Burt and Macdonald were performed [11, 12, 21]. With a sample size of 80 participants (40 females and 40 males) the 95% confidence intervals for the 2.5 and 97.5-percentile would be 1.4 times the standard deviation, which was considered an acceptable precision.

Kappa Cohen's statistic was used for inter-rater agreement in terms of quality grading (categories 1-3 vs. 4-5).

Baseline characteristics are presented by absolute and relative frequencies or by mean with standard deviation (sd) and median with lower and upper quartiles (lq, uq) for categorical and continuous variables, respectively. They were compared between women and men using Fisher's exact test for categorical variables, and Student's test and the Wilcoxon-Mann-Whitney test for and continuous variables.

The 2.5th, 25th, 50th, 75th and 97.5th percentiles were determined using quantile regressions with Stata command bsqreg. Standard errors were estimated via bootstrapping with 1000 repetition, confidence intervals and p-values were based on the t-distribution.

Correlations were assessed by Pearson product moment correlation and Spearman's rank-order correlation with 95% confidence intervals based on Fisher's transformation.

Simple linear regressions were fitted for radius and tibia strength including one covariate of interest, gender and their interaction. Marginal effects for women and men were calculated using Stata post-estimation command margins. Continuous predictors were standardized for a better comparability by subtracting the mean and dividing by the standard deviation and all effects are presented as change in kN of strength for a change of the predictor variable by one standard deviation. A p-value for interaction was calculated from a likelihood ratio test comparing models with and without the interaction term.

In addition, two multivariable linear regression models with and without total hip aBMD were fitted, and a pre-specified set of covariates (i.e. gait speed, hand grip dominant hand, dietary calcium, gender, age, weight, height). The two models were compared using likelihood ratio tests.

Basic assumptions of the linear regressions such as homoscedasticity and normality of the errors were checked visually using residuals-vs-fitted and quantile-quantile plots.

All statistical analyses were performed with Stata version 15 [StataCorp. 2017. *Stata Statistical Software: Release 15*. College Station, TX: StataCorp LLC.].

Results

Overall, 46 women and 41 men were recruited. Demographics including results of the clinical and DXA assessments are shown in Table 1. Baseline characteristics between men and women were not significantly different in terms of age, previous and parenteral fractures, and total hip or femoral neck aBMD. However, female and male participants differed significantly regarding weight and height, gait speed, hand grip, smoking status, and dietary calcium.

HR-pQCT

Results of HR-pQCT measurements are displayed in Table 2. Due to motion artefacts, radius images had to be repeated once in 18 men (44%) and 28 women (61%; *p*=0.11), and repeated twice in 3 men (7%) and 3 women (7%; *p*=0.88). Tibia images had to be repeated once in 12 men (29%) and 11 women (24%; *p*=0.57), and twice in 1 man (2%) and 2 women (4%; *p*=1.0). Overall, we had to exclude tibia exams of two women and one radius exam in one woman for analysis due to motion artefacts after three trials each. Significant differences were found between men and women for all HR-pQCT parameters. Men had a significantly higher total vBMD, trabecular vBMD, cortical perimeter and cortical thickness for both radius and tibia. In contrast, women had a significantly higher cortical vBMD and a significantly lower cortical porosity of both radius and tibia. Non-parametric statistics are shown in the appendix.

In hFE analysis, men had significant higher strength for radius than women (mean (sd) 6642 ± 1797 N vs. 4107 ± 1199 N; p<0.001) and tibia (18200 ± 4219 N vs. 11971 ± 3153 N; p<0.001). Similarly, stiffness was higher in men for radius (43443 ± 9293 N/mm vs. 29004 N/mm ± 6558 N; p<0.001) and tibia compared to women (70340 ± 14871 N/mm vs. 48320 ± 11034 N/mm; p<0.001) (Table 2).

Figure 2 displays the different percentiles with 95% confidence intervals (CI) of radius and tibia strength for women and men estimated by quantile regression. Absolute numbers thereof can be found in the appendix. As a sensitivity analysis, we performed age-dependent analyses in the subgroup of men. Assuming linear effects, radius and tibia strength decreased in men by -0.09 (95% CI -0.20 to 0.02, p=0.11) and -0.37 (95%CI -0.60 to -0.13, p=0.003) kN per year, respectively. Using piecewise linear effects with a knot at 30 years, there was almost no change up to 30 years (0.02 (95% CI -0.16 to 0.20, p=0.83) for radius strength and 0.01 (95% CI -0.37 to 0.38, p=0.98) for tibia strength) but a tendency for a decrease after 30 years (-0.36 (95% CI -0.75 to 0.02, p=0.07) for radius and -1.32 (95% CI -2.12 to -0.52, p=0.002) for tibia).

Correlation Analyses

Radius and tibia strength were highly and significantly correlated with stiffness for both women (Pearson correlation coefficient radius 0.97 (95% CI 0.95-0.98), p-value <0.001; tibia 0.99 (0.98-0.99), p<0.001) and men (radius 0.97 (0.95-0.99), p<0.001; tibia 0.99 (0.98-0.99), p<0.001).

Total hip aBMD by DXA correlated significantly with strength at radius and tibia for both women (Pearson correlation coefficient radius 0.30 (95% CI 0.00-0.54), p=0.048); tibia 0.57 (0.33-0.74), p<0.001) and men (radius 0.59 (0.35-0.76), p<0.001; tibia 0.75 (0.57-0.86), p<0.001).

Simple Regression Analyses

In simple regression models, an increase in femoral neck aBMD by one standard deviation (sd) lead to a significant increase of 0.72 kN (95% CI 0.30-1.14) for radius strength in women and 1.09 kN (0.72-1.45) for radius strength in men (Appendix). Similarly, increase of femoral neck aBMD by one sd lead to an increase of 2.52 kN (1.67-3.37) in women for tibia strength and 3.17 kN (2.44-3.90) in men for tibia strength. Besides, age was only significantly associated with tibia strength in men. We did not find any evidence, that other variables including gait speed, hand grip, dietary calcium, height and weight were predictive for radius or tibia strength. Neither did we find evidence that any of the variables would have different effects in men and women.

Multivariable Regression Analyses

In the two multivariable linear regression models with and without femoral neck aBMD, the addition of femoral neck aBMD significantly improved the model (p<0.001). In the multivariable model including femoral neck aBMD, both femoral neck aBMD and gender were predictive for both radius and tibia strength (Figure 3). In the multivariable model without femoral neck aBMD, female gender was predictive for significant lower radius strength, and increasing age and female gender were predictive for significant lower tibia strength, respectively (Appendix).

Discussion

In this study, we present reference values of microstructural bone parameters and finite element analysis for HR-pQCT in the young healthy male and female Swiss population.

To our knowledge, this is the first study combining the approach of multiple stack HR-pQCT (XCT II, SCANCO Medical AG) with hFE analysis for estimation of compressive bone strength. The methodology of multiple stack measurements features multiple advantages. First, the weakest part of the bone for the assessment of bone strength is included in the region of interest. Second, the calculation of bone strength is less dependent on the applied boundary conditions in the finite element simulations. Third, the region of interest that can be registered in longitudinal studies is about twice as large. Finally, a patient size-dependent region of interest can be defined during post-processing instead during the clinical examination. Two disadvantages of multiple stacks are the extended scanning time and the additional overall radiation dose. The latter is intrinsically low, while the former is currently dominated by other factors such as patient installation and scanning software operations. Despite the above arguments, the cost/benefit analysis of multiple stacks remains to be established on clinical grounds. Regarding the finite element approach, Arias-Moreno et al. [19] recently demonstrated, that it is possible to predict consistently stiffness and strength using homogenized-FE (hFE) approach in comparison to micro-FE (µFE). The major advantage of hFE approach lies in the fact, that hFE protocols are computed much faster than µFE predestinating hFE approach for clinical use. The *ex vivo* reproducibility of the strength estimation by hFE was reported in Hosseini et al. [4], while the *in vivo* reproducibility is currently under investigation.

When comparing our results to those of prior evidence we have to be aware of the fact, that absolute numbers cannot be directly compared, because major methodological approaches differ substantially between studies (hFE vs. µFE, single stack vs. multiple stack, HR-pQCT device). To put our main outcome of strength into the context of recent evidence, we summarized results of selected population-based evidence of failure load/strength [10-12, 21-25] in Table 3. Both absolute values of radius and tibia strength are about twice as high as in Burt et al. (2014) [12] and (2016) [21]. Thus, relative mean differences of radius and tibia strength between data by Burt et al. and our study are similar.

These differences of absolute values of strength may not be explained by the fact, that we used double resp. triple stacks in combination with hFE analysis, since the weakest link theory suggests that multiple stacks should deliver lower strength than single stacks. However, the observed differences might be explained by the fact, that we used non-linear hFE analysis (with equal accuracy [4]) instead of the

current linear µFE state-of the art. In fact, prior studies using single stack approach and µFE used different biomechanical testing protocols for their calibration [26]. The material parameters employed in the current hFE method rely on the biomechanical tests developed by Hosseini et al. [4] and were calibrated by Arias-Moreno et al. [19].

Regarding microstructural parameters of HR-pQCT our results are in accordance to data of the CAMOS cohort [21] both showing significantly higher values in men than in women for nearly all. In accordance to our results CAMOS describes a higher cortical vBMD for women compared to men in tibia, but not in radius. Our results, however, suggest significant higher indices of cortical vBMD in women compared to men for both radius and tibia. In line with earlier findings we observed men having thicker, more dense bones in comparison to women resulting higher bone strength and stiffness at both radius and tibia. However, men feature less dense and more porous cortices. Various reasons for these sex differences have been discussed e.g. later pubertal growth, bigger bone size, higher levels of calcium intake, and physical activity in men compared to women [27-29].

In an explorative way, we tried to identify clinical predictors of compressive bone strength. In the multivariable model without DXA femoral neck aBMD we found the already known factors of age and gender to be significant predictors of bone strength. Interestingly, we did not find evidence, that variables of clinical assessment (hand grip, gait speed, dietary calcium) were predictive for bone strength. Nevertheless, due to limited sample size, we cannot exclude an effect of clinical parameters that was missed hereby. As we specifically looked for linear relationships, we cannot exclude that non-linear relationships might be present. In the multivariable model including femoral neck aBMD, the latter was the strongest independent predictor of bone strength showing stronger association with tibia strength than radius strength. Because both femur (femoral neck aBMD) and tibia (HR-pQCT) are weight-bearing bones of the lower extremity this stronger association is easily comprehensible and is consistent with the observation that tibial strength measured by HR-pQCT discriminates incident fractures better than radius strength [8]. Thus, this observation could have implications for future clinical practice e.g. favoring tibia examination over radius upon HR-pQCT exams. The fact that when adding aBMD to the multivariable regression model, age dropped out to be an independent predictor, reflects that the morphology-mechanical property relationships are invariant with age in the narrow age range of this young population.

There are several limitations to our study. The study was performed in Caucasian Swiss young people, featuring limited variation in age. Therefore, our results are considered representative for a Caucasian population, but may not be applied to other populations and may not be representative for other cultural contexts. Second, multivariable models have to be interpreted with caution due to limited

sample size. We computed these models to identify predictors in an explorative way. Therefore, we predefined two models including a maximum of 9 variables to account for the limited sample size. Finally, we did not perform direct comparison of hFE with µFE and multiple with single stack analyses, respectively. The predefined nature of this project was not on methodological purposes, but the focus of this study was to establish a standardized clinically feasible approach providing the basis for clinical implementation.

So far, HR-pQCT was almost exclusively used as a research tool and was not used in clinical practice. Before implementation into clinical practice can take place, several feasibility criteria have to be met. Technical, economic, and operational factors play an important role when coming to review feasibility of the HR-pQCT. To best meet clinical expectations, we used the most validated, reproducible and clinically feasible approach in our methodology of performing HR-pQCT. In comparison to DXA, HR-pQCT takes longer for scan data acquisition and computing. Although one HR-pQCT measurement takes 3 to 5 minutes of time for one scan, measurements often had to be repeated due to motion artefacts. Half of radius and one third of tibia measurements had to be repeated suggesting that chest movement due to breathing may induce nearby arm movements diminishing quality of radius images. However, substantial time saving was achieved by using hFE instead of µFE analysis. Charged CPU times for double section at the radius where about 20 minutes and for triple section at the tibia about 6 hours. Therefore, the time of computing is substantially reduced compared to µFE. This benefit may be improved further by restricting the hFE analysis to a single linear step delivering stiffness, but at the cost of the higher dependence of this indirect strength estimation on section thickness.

Overall, one of the foundations for clinical implementation has now been established by the presented data of our study. These results provide a basis to help develop a new clinical approach for predicting fracture risk.

In conclusion, reference values for radius and tibia strength using multiple stacks HR-pQCT with hFE analysis are presented enabling clinical implementation. Gender and femoral neck aBMD were independent predictors of bone strength.

References

1. Hauger AV, Bergland A, Holvik K, Stahle A, Emaus N, Strand BH (2018) Osteoporosis and osteopenia in the distal forearm predict all-cause mortality independent of grip strength: 22-year follow-up in the population-based Tromso Study. Osteoporosis international : a journal established as result of cooperation between the European Foundation for Osteoporosis and the National Osteoporosis Foundation of the USA 29:2447-2456
2. Stone KL, Seeley DG, Lui LY, Cauley JA, Ensrud K, Browner WS, Nevitt MC, Cummings SR, Osteoporotic Fractures Research G (2003) BMD at multiple sites and risk of fracture of multiple types: long-term results from the Study of Osteoporotic Fractures. Journal of bone and mineral research : the official journal of the American Society for Bone and Mineral Research 18:1947-1954
3. Siris ES, Brenneman SK, Barrett-Connor E, Miller PD, Sajjan S, Berger ML, Chen YT (2006) The effect of age and bone mineral density on the absolute, excess, and relative risk of fracture in postmenopausal women aged 50-99: results from the National Osteoporosis Risk Assessment (NORA). Osteoporosis international : a journal established as result of cooperation between the European Foundation for Osteoporosis and the National Osteoporosis Foundation of the USA 17:565-574
4. Hosseini HS, Dunki A, Fabech J, Stauber M, Vilayphiou N, Pahr D, Pretterklieber M, Wandel J, Rietbergen BV, Zysset PK (2017) Fast estimation of Colles' fracture load of the distal section of the radius by homogenized finite element analysis based on HR-pQCT. Bone 97:65-75
5. Nishiyama KK, Macdonald HM, Hanley DA, Boyd SK (2013) Women with previous fragility fractures can be classified based on bone microarchitecture and finite element analysis measured with HR-pQCT. Osteoporosis international : a journal established as result of cooperation between the European Foundation for Osteoporosis and the National Osteoporosis Foundation of the USA 24:1733-1740
6. Stein EM, Liu XS, Nickolas TL, et al. (2010) Abnormal microarchitecture and reduced stiffness at the radius and tibia in postmenopausal women with fractures. Journal of bone and mineral research : the official journal of the American Society for Bone and Mineral Research 25:2572-2581
7. Wang J, Stein EM, Zhou B, Nishiyama KK, Yu YE, Shane E, Guo XE (2016) Deterioration of trabecular plate-rod and cortical microarchitecture and reduced bone stiffness at distal radius and tibia in postmenopausal women with vertebral fractures. Bone 88:39-46
8. Samelson EJ, Broe KE, Xu H, et al. (2019) Cortical and trabecular bone microarchitecture as an independent predictor of incident fracture risk in older women and men in the Bone Microarchitecture International Consortium (BoMIC): a prospective study. The lancet Diabetes & endocrinology 7:34-43

9. Kroker A, Plett R, Nishiyama KK, McErlain DD, Sandino C, Boyd SK (2017) Distal skeletal tibia assessed by HR-pQCT is highly correlated with femoral and lumbar vertebra failure loads. Journal of biomechanics 59:43-49

10. Khosla S, Riggs BL, Atkinson EJ, Oberg AL, McDaniel LJ, Holets M, Peterson JM, Melton LJ, 3rd (2006) Effects of sex and age on bone microstructure at the ultradistal radius: a population-based noninvasive in vivo assessment. Journal of bone and mineral research : the official journal of the American Society for Bone and Mineral Research 21:124-131

11. Macdonald HM, Nishiyama KK, Kang J, Hanley DA, Boyd SK (2011) Age-related patterns of trabecular and cortical bone loss differ between sexes and skeletal sites: a population-based HR-pQCT study. Journal of bone and mineral research : the official journal of the American Society for Bone and Mineral Research 26:50-62

12. Burt LA, Macdonald HM, Hanley DA, Boyd SK (2014) Bone microarchitecture and strength of the radius and tibia in a reference population of young adults: an HR-pQCT study. Archives of osteoporosis 9:183

13. Baumbach SF, Schmidt R, Varga P, Heinz T, Vecsei V, Zysset PK (2011) Where is the distal fracture line location of dorsally displaced distal radius fractures? Journal of orthopaedic research : official publication of the Orthopaedic Research Society 29:489-494

14. Mueller TL, van Lenthe GH, Stauber M, Gratzke C, Eckstein F, Muller R (2009) Regional, age and gender differences in architectural measures of bone quality and their correlation to bone mechanical competence in the human radius of an elderly population. Bone 45:882-891

15. Varga P, Pahr DH, Baumbach S, Zysset PK (2010) HR-pQCT based FE analysis of the most distal radius section provides an improved prediction of Colles' fracture load in vitro. Bone 47:982-988

16. Hars M, Audet MC, Herrmann F, De Chassey J, Rizzoli R, Reny JL, Gold G, Ferrari S, Trombetti A (2018) Functional Performances on Admission Predict In-Hospital Falls, Injurious Falls, and Fractures in Older Patients: A Prospective Study. Journal of bone and mineral research : the official journal of the American Society for Bone and Mineral Research 33:852-859

17. Bonaretti S, Majumdar S, Lang TF, Khosla S, Burghardt AJ (2017) The comparability of HR-pQCT bone measurements is improved by scanning anatomically standardized regions. Osteoporos Int

18. Pauchard Y, Liphardt AM, Macdonald HM, Hanley DA, Boyd SK (2012) Quality control for bone quality parameters affected by subject motion in high-resolution peripheral quantitative computed tomography. Bone 50:1304-1310

19. Arias-Moreno AJ, Hosseini HS, Bevers M, Ito K, Zysset P, van Rietbergen B (2019) Validation of distal radius failure load predictions by homogenized- and micro-finite element analyses based on second-generation high-resolution peripheral quantitative CT images. Osteoporosis international : a

journal established as result of cooperation between the European Foundation for Osteoporosis and the National Osteoporosis Foundation of the USA 30:1433-1443

20. Harrigan TP, Mann, R. W. Characterization of microstructural anisotropy in cancellous bone using a second rank tensor. Journal of Materials Science 19:761–767

21. Burt LA, Liang Z, Sajobi TT, Hanley DA, Boyd SK (2016) Sex- and Site-Specific Normative Data Curves for HR-pQCT. Journal of bone and mineral research : the official journal of the American Society for Bone and Mineral Research 31:2041-2047

22. Dalzell N, Kaptoge S, Morris N, Berthier A, Koller B, Braak L, van Rietbergen B, Reeve J (2009) Bone micro-architecture and determinants of strength in the radius and tibia: age-related changes in a population-based study of normal adults measured with high-resolution pQCT. Osteoporosis international : a journal established as result of cooperation between the European Foundation for Osteoporosis and the National Osteoporosis Foundation of the USA 20:1683-1694

23. Sode M, Burghardt AJ, Kazakia GJ, Link TM, Majumdar S (2010) Regional variations of gender-specific and age-related differences in trabecular bone structure of the distal radius and tibia. Bone 46:1652-1660

24. Hansen S, Shanbhogue V, Folkestad L, Nielsen MM, Brixen K (2014) Bone microarchitecture and estimated strength in 499 adult Danish women and men: a cross-sectional, population-based high-resolution peripheral quantitative computed tomographic study on peak bone structure. Calcified tissue international 94:269-281

25. Hung VW, Zhu TY, Cheung WH, Fong TN, Yu FW, Hung LK, Leung KS, Cheng JC, Lam TP, Qin L (2015) Age-related differences in volumetric bone mineral density, microarchitecture, and bone strength of distal radius and tibia in Chinese women: a high-resolution pQCT reference database study. Osteoporosis international : a journal established as result of cooperation between the European Foundation for Osteoporosis and the National Osteoporosis Foundation of the USA 26:1691-1703

26. Macneil JA, Boyd SK (2008) Bone strength at the distal radius can be estimated from high-resolution peripheral quantitative computed tomography and the finite element method. Bone 42:1203-1213

27. Gordon CL, Halton JM, Atkinson SA, Webber CE (1991) The contributions of growth and puberty to peak bone mass. Growth, development, and aging : GDA 55:257-262

28. Seeman E (2001) Clinical review 137: Sexual dimorphism in skeletal size, density, and strength. The Journal of clinical endocrinology and metabolism 86:4576-4584

29. Burrows M, Liu D, Moore S, McKay H (2010) Bone microstructure at the distal tibia provides a strength advantage to males in late puberty: an HR-pQCT study. Journal of bone and mineral research : the official journal of the American Society for Bone and Mineral Research 25:1423-1432

Tables

Table 1. Baseline Characteristics		**Women (n=46)**	**Men (n=41)**	***p*-value***
Demographics				
Age (years),	mean (SD)	25.1 (5.0)	26.2 (5.2)	0.32
	median (lq, uq)	23.5 [21.0, 28.0]	25.0 [22.0, 31.0]	0.27
Weight (kg)	mean (SD)	62.5 (9.3)	77.1 (12.4)	<0.001
	median (lq, uq)	61.0 [57.0, 66.0]	74.0 [70.0, 81.0]	<0.001
Height (cm)	mean (SD)	167 (5.3)	180 (6.4)	<0.001
	median (lq, uq)	166.0 [163.0, 170.0]	179.0 [175, 184]	<0.001
BMI (kg/m^2)	mean (SD)	22.5 (2.9)	23.9 (3.4)	0.04
	median (lq, uq)	22.2 [20.6, 23.4]	23.1 [22.2, 25.9]	0.05
Clinical risk factors				
Previous low trauma fracture (n)				
	mean (SD)	0	0	1.00
	median (lq, uq)	0	0	1.00
Parental hip fracture (n), mean (SD)		1 (2%)	0 (0%)	0.50
	median (lq, uq)	1 (2%)	0 (0%)	1.00
Dietary calcium (mg)		679(261)	871 (309)	0.002
	median (lq, uq)	619 [520, 810]	850 [624, 1130]	0.003
Physical assessment				
Gait speed (m/sec), mean (SD)		0.70 (0.09)	0.64 (0.13)	0.02
	median (lq, uq)	0.7 [0.6, 0.7]	0.6 [0.6, 0.7]	<0.001
Repeated chair rise test (sec)		7.10 (1.29)	7.12 (1.43)	0.84
	median (lq, uq)	7.1 [6.0, 8.1]	7.1 [6.0, 8.1]	0.94
Hand grip dominant hand (kg)		28.1 (4.6)	44.9 (8.2)	<0.001
	median (lq, uq)	28.9 [24.0, 30.6]	45.0 [39.3, 51.3]	<0.001
DXA				
Total hip aBMD (g/cm^2), mean (SD)		0.95 (0.13)	1.06 (0.16)	<0.001
	median (lq, uq)	0.94 [0.85, 1.06]	1.06 [0.93, 1.19]	0.001
Total hip t-score (-), mean (SD)		0.08 (1.07)	0.26 (1.08)	0.42
	median (lq, uq)	-0.05 [-0.70, 1.00]	0.20 [-0.60, 1.00]	0.54
Total hip z-score (-),mean (SD)		0.09 (1.06)	0.30 (1.07)	0.36
	median (lq, uq)	-0.05 [-0.70, 1.00]	0.20 [-0.60, 1.10]	0.45
Neck hip aBMD (g/cm^2), mean (SD)		0.83 (0.12)	0.92 (0.15)	0.004

median (lq, uq)	0.82 [0.76, 0.94]	0.92 [0.80, 1.02]	0.008
Neck hip t-score (-),mean (SD)	-0.15 (1.09)	-0.03 (1.12)	0.60
median (lq, uq)	-0.30 [-0.80, 0.80]	-0.10 [-1.00, 0.70]	0.78
Neck hip z-score (-),mean (SD)	-0.11 (1.05)	0.08 (1.06)	0.42
median (lq, uq)	-0.30 [-0.60, 0.80]	-0.10 [-0.80, 0.90]	0.59

*From Fisher's exact test for categorical variable, Student's t-test for continuous variables presented with mean (sd) and the Wilcoxon-Mann-Whitney test for continuous variables presented with median [lq, uq]

†Dietary calcium missing for 1 women.

Abbreviations: sd, standard deviation; lq, lower quartile, uq, upper quartile, n.a., not applicable; aBMD, areal bone mass density.

Table 2. Summary of HR-pQCT and hFE Outcomes at the Distal Radius and Distal Tibia for Women and Men

	Women		Men	
	Mean (SD)	**95% CI**	**Mean (SD)**	**95% CI**
Radius				
Left side (n)	38		34	
Good Quality of scans, n (%)	45 (97.8%)		41 (100.0%)	
Total vBMD* (mg HA/ccm)	298 (49.4)	283 - 313	335 (52.0)	318 - 351
Trabecular vBMD* (mg HA/ccm)	145 (32.2)	135 - 154	200 (36.5)	188 - 211
Cortical vBMD* (mg HA/ccm)	931 (37.6)	919 - 942	890 (32.2)	879 - 900
Cortical perimeter* (mm)	65.8 (4.6)	64.4 - 67.1	75.8 (5.3)	74.1 - 77.5
Cortical porosity* (-)	0.003 (0.002)	0.003 - 0.004	0.007 (0.006)	0.005 - 0.009
Cortical thickness* (mm)	1.06 (0.15)	1.01 - 1.11	1.21 (0.21)	1.15 - 1.28
Trabecular area* (mm^2)	226 (36.9)	215 - 237	290 (51.9)	274 - 306
Cortical area* (mm^2)	55.2 (7.1)	53.1 - 57.4	71.3 (10.2)	68.1 - 74.6
Strength* (N)	4110 (1200)	3747 - 4468	6640 (1800)	6075 - 7209
Stiffness* (N/mm)	29000 (6560)	27033 - 30974	43400 (9290)	40510 - 46376
Tibia				
Left side (n)	38		34	
Good quality of scans, n (%)	44 (95.7%)		41 (100.0%)	
Total vBMD† (mg HA/ccm)	266 (36.7)	254 - 277	301 (43.3)	287 - 314
Trabecular vBMD† (mg HA/ccm)	197 (31.7)	187 - 206	238 (33.5)	227 - 248
Cortical vBMD† (mg HA/ccm)	866 (30.5)	857 - 875	820 (37.0)	808 - 832
Cortical perimeter† (mm)	113 (6.9)	111 - 116	126 (8.6)	123 - 129
Cortical porosity† (-)	0.015 (0.006)	0.013 - 0.017	0.027 (0.010)	0.023 - 0.030
Cortical thickness† (mm)	1.09 (0.16)	1.04 - 1.13	1.31 (0.30)	1.22 - 1.40
Trabecular area† (mm^2)	814 (105)	782 - 846	993 (147)	947 - 1040
Cortical area† (mm^2)	95.2 (12.8)	91.3 - 99.1	121 (24.3)	114 - 129
Strength† (N)	11971 (3150)	11013 - 12930	18200 (4220)	16868 - 19531
Stiffness† (N/mm)	48300 (11000)	44965 - 51675	70340 (14900)	65646 - 75034

*Missing for one woman; †Missing for two women.

Abbreviations: HR-pQCT, high resolution pQCT; hFE, homogenized finite element analysis; vBMD, volumetric bone mineral density; sd, standard deviation; lq, lower quartile; uq, upper quartile; CI, confidence interval.

Statistical significance between male and female with p≤0.001 except for left side and good quality.

Table 3. Summary of Selected Population-based Evidence on Estimates of Failure Load/Strength

	Strength/Failure load (N)					
	Radius			Tibia		
Author (Year of Publication)	Men mean (SD)	Women mean (SD)	Absolute and relative mean difference men-women (%)	Men mean (SD)	Women mean (SD)	Absolute and relative mean difference men-women (%)
Khosla (2006) [10]	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
Dalzell (2009) [22]	2789 (49)	n.a.	n.a.	2686 (49)	n.a.	n.a.
Sode (2010) [23]	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
Hansen (2014) [24]	5920 (822)	3993 (731)	1927 (48%)	15054 (191)	10923 (1721)	4153 (38%)
Burt (2014)[b)] [12]	3300.6 (570.4)	2068.3 (323.0)	1232.3 (60%)	8059.8 (1469.3)	5654.1 (891.1)	2405.7 (43%)
Hung (2015) [25]	n.a.	2604 (483)	n.a.	n.a.	6282 (1160)	n.a.
Burt (2016)[a)] [21]	3073.9 (n.a.)	2133.0 (n.a.)	940.9 (44%)	7894.1 (n.a.)	6073.8 (n.a.)	1820.3 (30%)
Stuck (2019)[d)]	**6642 (1797)**	**4107 (1199)**	**2535 (62%)**	**18200 (4219)**	**11971 (3153)**	**6229 (52%)**

Abbreviations: SD, standard deviation; n.a., not assessed

a) 50th percentile of "age 20's" displayed

b) Mean (SD) of age group "25-29 years" are displayed

c) Median (interquartile range) displayed according to paper

d) Present study

Figure Legends

Figure 1 (Panel A). Reference line and extension of the two stacks covering Colles' fracture zone.

Figure 1 (Panel B). Reference line and positions of the three stacks covering the epiphysis and metaphysis of the distal tibia.

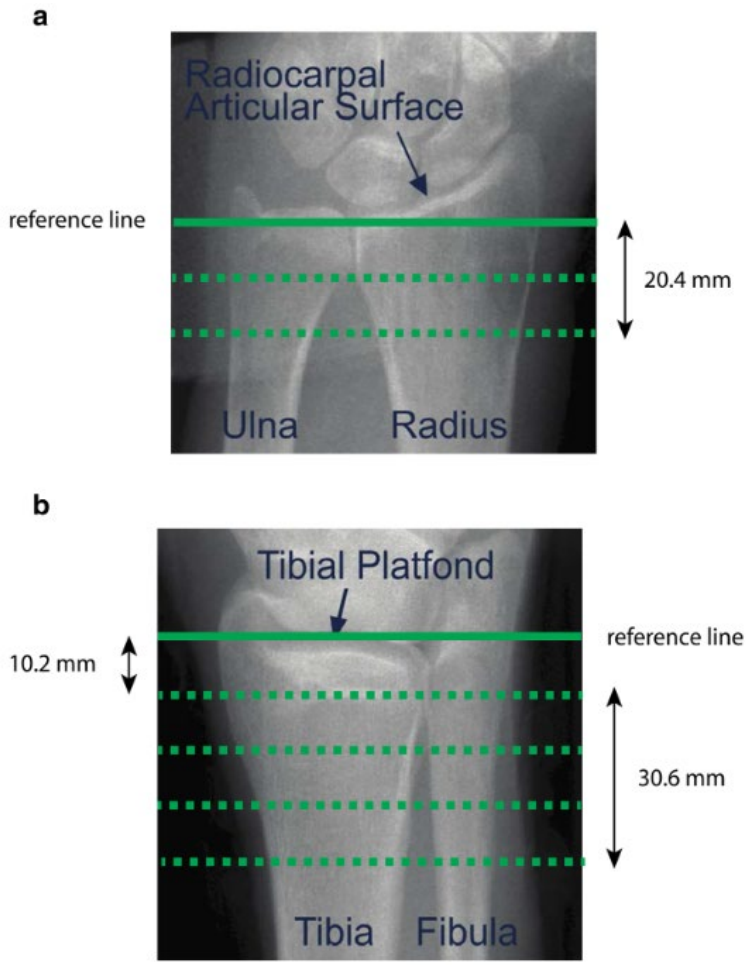

Figure 2. Percentiles of radius and tibia strength for women and men estimated by quantile regression. The raw data is indicated with circles. Age was not included in the estimation of the quantiles.

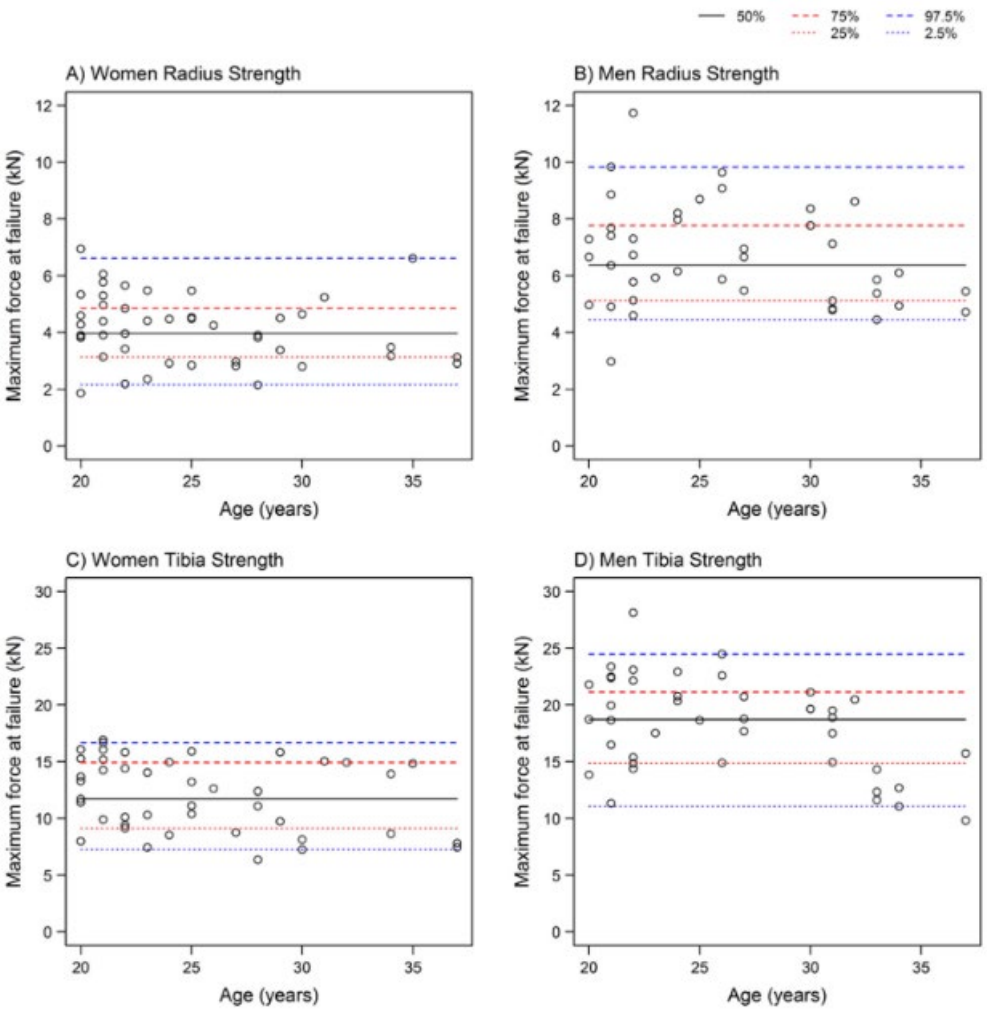

Figure 3. Multivariable linear regression model for radius and tibia strength including total hip aBMD.

Abbreviations: aBMD, areal bone mineral density; sd, standard deviation.